LEVERAGING MACHINE LEARNING, DEEP LEARNING, AND NATURAL LANGUAGE PROCESSING TECHNIQUES TO COMBAT FAKE NEWS IN ONLINE SOCIAL NETWORKS

A SYNOPSIS

Submitted by

S. SELVA BIRUNDA

(Reg. No. 201907175)

in partial fulfillment for the award of the degree of

DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SCHOOL OF COMPUTING

KALASALINGAM ACADEMY OF RESEARCH AND EDUCATION

(DEEMED TO BE UNIVERSITY)

ANAND NAGAR, KRISHNANKOIL – 626 126

TAMIL NADU, INDIA

JANUARY 2024

1. INTRODUCTION

Online Social Network is a platform where people can connect and communicate with each other online. Some Online Social Networks include Facebook, Instagram, Twitter, and WhatsApp. People can share their views and opinions, and connect with like-minded people easily through these networks. Anyone can post any information that they like. Now, a question arises can people believe and rely on these Online Social Networks for news information? The answer to this question is definitely not because every day millions and trillions of information get propagated. There can be no assurance that all the information propagated via Online Social Networks is real. People trust these Online Social Networks blindly without evaluating the veracity and origin of the information. In a nutshell, one of the issues to be addressed in Online Social Networks in recent days is Fake news. Fake news is not new as it existed many years before, however, the term was coined and became a public concern during the 2016 U.S. presidential election [1]. Fake news is junk news or pseudo-news shared intentionally or by manipulating the real information. Fake news can take various forms namely, text, hyperlinks, multimedia, video, audio, and images. This research work focuses only on the text content. Fake news impacts almost all fields namely, politics, health care, media, and many more. Recently, during the COVID-19 pandemic era, misinformation about COVID-19 treatments, vaccines, and prescriptions has arisen in vast amounts [2]. Misinformation regarding the COVID-19 pandemic is mentioned as a massive “infodemic”. During that time, Online social networks, especially, Twitter is utilized 8.4% more than the previous times, and a vast amount of misinformation was propagated through the platform [3]. A lot of research has been going on in the aspect of mitigating fake information, however, misinformation spread has not been stopped or reduced to some extreme.

Some of the challenges related to COVID-19 Fake tweet detection include the lack of a “Ground Truth” Tweet dataset,

annotation inconsistency, inaccurate supervision, speed of misinformation dissemination, limited space for source credibility verification, and the need for real-time detection of COVID-19 tweets [4]. Hence, the Fake Tweets Detection model that addresses the above challenge is essential because, still there is no fake news detection system employed that can completely detect fake news with maximum accuracy.

Optimal feature selection-based fake news detection considers the unstructured tweet dataset as a searching space and metaheuristic algorithms as a search method for the COVID-19 fake tweet detection problem. Some of the common challenges in this domain include computational complexity, premature convergence, and lack of convergence to Global optima [5]. Hence, the optimal feature selection-based fake news detection model that addresses these challenges is essential.

As misinformation arises day by day, timely detection is necessary to combat the impact of that information. Besides, the speed of dissemination of misinformation is taken into account, and very limited data is available for fake news detection, which is not sufficient for the supervised learning approach [6]. Developing a self-trained COVID-19 fake news detection system without latency is still challenging.

Hence, designing interventions to address these issues can be aided by research in these fields.

2. LITERATURE REVIEW

2.1 Fake Tweets Detection

Detection of fake news by cross-verifying it with reliable sources using K-means clustering was done by Ghadiri et al. [7]. Reliable sources of information were used as a reference to check the reliability of the news. The detection happened with less accuracy because of the news clustering issue. Clustering of news was done

based on the vocabulary and the vocabulary of one tweet matches the multiple clusters. In addition, optimized source features for short tweets were not considered. Detecting the fake tweets by extracting network, linguistic-related features, and SMOTE over-sampling technique was proposed by Nikiforos et al. [8]. Grid searching was applied for tuning the hyperparameters, however, it took more computation time. The authors of Qureshi et al. [9] detected the COVID-19 fake news by employing Machine Learning algorithms, namely, SVM, Linear Regression, Decision tree, and Gradient Boosting algorithm. The technique employed for choosing the parameters of the Gradient Boosting algorithm was not defined. Consequently, the challenges faced by the existing fake news detection model by applying the Gradient Boosting algorithm include computational complexity, lack of Performance variability resulting from the choice of specific hyperparameters (need for fine-tuning) Capuano et al. [10], and potential overfitting.

2.2 Optimal Feature Selection

The authors of Ozbay & Alatas [11] utilized Grey Wolf Optimization (GWO) and Salp Swarm optimization (SSO) for the detection of fake news. GWO had the challenge of a slow convergence rate. The authors of Abd Elaziz et al. [12] applied the Fire Hawk optimizer algorithm to extract the optimal features from Arabic fake news. Further, to improve the performance, opposition-based optimization techniques could be utilized. The Firefly Optimization algorithm was applied to detect fake news by tuning the parameters namely, the size of the population, distance measure to calculate attraction, and local search strategy strength Jaiswal et al. [13]. The Hunter Prey Optimization (HPO) algorithm was developed for the optimal hyperparameter tuning of the Long Short Term Memory (LSTM)-Recurrent Neural Network (RNN) model Alshahrani et al. [14]. This could be enhanced by the feature selection process. The authors of Nithya & Sahayadhas [15] applied a Hybrid Squirrel Dragonfly search optimization algorithm for fake news detection to

solve the issues of minimum exploitation and low convergence rate in local optima obtained by SSO.

2.3 Credibility detection of emerging tweets

The authors of Patwa et al. [16] detected the COVID-19 fake news by applying supervised algorithms namely, Logistic Regression, Support Vector Machine, Decision Tree, and Gradient Boosting. The supervised approach required more labeled training data and annotation was done manually. The misleading information on COVID-19 was detected using supervised Machine Learning algorithms, however, they failed to detect the misleading information in real time Elhadad et al. [17]. Besides, the dataset used here is not a ground truth dataset, hence, it could be enriched by adding information from Twitter's official handles namely, the World Health Organization, Centers for Disease Control of Disease and Prevention, and International Committee of the Red Cross. The authors of Paka et al. [18] developed a cross-stitch semi-supervised neural attention model for the detection of COVID-19 fake news in a semi-supervised approach, however, could not detect the fake news without latency. Consequently, the challenges faced by the existing semi-supervised models include annotation inconsistency, memory-intensive issues, lack of self-trained algorithms and early detection of fake news is not possible.

The proposed research work addresses these challenges and contributes an efficient Fake News Detection model.

3. RESEARCH PROBLEM

Based on the conducted research, the following issues consistently emerge as usual and persistent challenges within the realm of COVID-19 fake news:

- Most of the existing fake news detection models are limited to content features and fail to consider the source's credibility. Besides the detection could be enhanced by fine-tuning the hyperparameters and checking the performance deviations.

- Most of the existing meta-heuristic optimization feature selection fake news detection models suffer from low convergence rates, trapped at local optima. In particular, the existing Honey Badger metaheuristic algorithm suffers from being trapped in local optima while performing the optimization process because of a loss in population diversity.
- The existing semi-supervised fake news detection models suffer from annotation inconsistency, memory-intensive issues, and detection of fake news without latency.

4. RESEARCH OBJECTIVE

Our major research emphasis is to develop a robust and efficient fake news detection method to maintain the integrity of information dissemination in Online Social Networks. This leads us to outline the following as our research objectives.

1. To automate the accurate detection of fake tweets and enhance the space for source verification by developing a Gradient-Boosting algorithm with hyperparameter tuning
2. To develop a novel Light Weight Convolutional Random Forest based Honey Badger optimization algorithm for optimal feature selection in COVID-19 fake tweet detection
3. To automatically detect the reliability of emerging tweets without delay by developing a novel Self-Trained Semi-Supervised Hybrid Deep Learning model

5. CONTRIBUTIONS

5.1 Source Credibility and Content-based COVID-19 Fake Tweets Detection Model

Source credibility and content-based fake tweet detection model is developed to automate the accurate detection of fake tweets that deviate from reliable sources (i.e. Twitter handles). The architecture of the proposed source credibility and content-based COVID-19 fake tweets detection model is depicted in Figure 1.

Initially, the tweet dataset curation happens by developing a snscrape script to extract the tweets from Twitter with the help of a keyword list. The curated dataset undergoes preprocessing steps namely, tokenization, stop words removal, punctuation removal, normalization, and lemmatization.

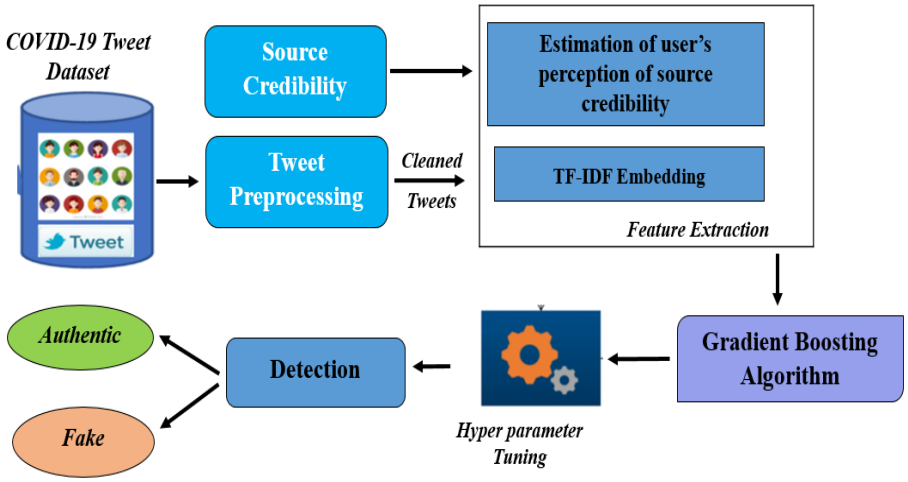

Figure 1: Proposed Architecture of source credibility and content-based COVID-19 fake tweets detection model

Now, the cleaned tweets are given as input to the Term Frequency-Inverted Document Frequency (TF-IDF) technique to generate vector representations for each word in the tweet. TF-IDF is employed to assess the term's significance within the tweet as well as its relative rarity throughout the entire dataset. The source credibility features are extracted based on the estimation of the user's perception. The extracted source credibility features as well as the tweet content features have complicated nonlinear synergy between them, therefore, to capture these synergies, an ensemble Gradient Boosting algorithm is developed.

5.1.1 Hyperparameter Tuned Gradient Boosting Algorithm

This algorithm combines multiple weak learner models to become strong learner models to make predictions. Subsampling is employed to reduce the computational complexity. The size of the

subset is fixed by the hyperparameter namely, “subsample”. Now, the models get trained iteratively and sequentially using the subset of features to overcome the errors obtained by the previous models. To obtain the optimal combination of hyperparameters in less time, the Random search technique is incorporated. The potential overfitting issue is addressed by adding regularization techniques namely, learning rate and tree depth. By specifying the learning rate ranges, the value of the learning rate is fixed by the random search. After fine-tuning the parameters of the proposed model by the Random search technique, the resulting model is tested to identify COVID-19 fake tweets. Finally, the challenges faced by the existing fake news detection model are addressed by the proposed model.

5.1.2 Experimental Analysis

Experimentation of the proposed model is carried out using Python and Scikit Learn library. The performance of the proposed model is analyzed using different experimental measures namely, accuracy, precision, recall, and F measure. The proposed model was compared with the Machine Learning classifiers namely, Support Vector Machine, K-Nearest Neighbors, Multinomial Naive Bayes (NB), Logistic Regression, Random Forest, AdaBoost, Decision Tree, and Gradient Boosting. The accuracy comparison of multiple classifiers is depicted in Figure 2.

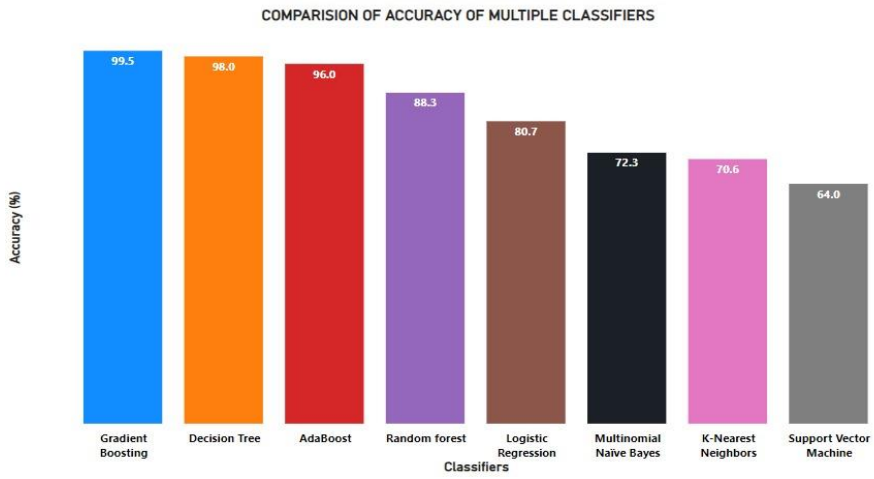

Figure 2: Accuracy Comparison of Multiple Classifiers

Hyperparameter-tuned Gradient Boosting achieves 99.5% more in detecting COVID-19 fake tweet detection than all other classifiers. The performance estimation of aforementioned classifiers in terms of precision, recall, and F1 Score is shown in Figure.3. The proposed algorithm achieves high precision, recall, and F1 Score of 0.99, 0.99, and 0.99 respectively.

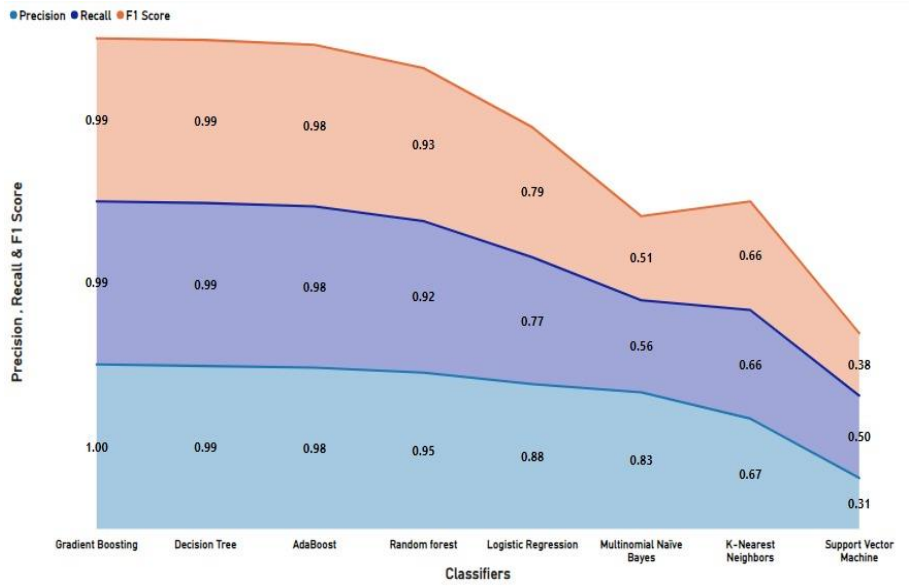

Figure 3: Performance comparison in terms of precision, Recall, and F1 score

5.2 Optimal Feature Selection Based COVID-19 Fake Tweets Detection

A novel Light Weight Convolutional Random Forest-based Honey Badger Optimization Algorithm (LCRF-HB) is proposed for optimal feature selection. The proposed architecture is depicted in Figure 4.

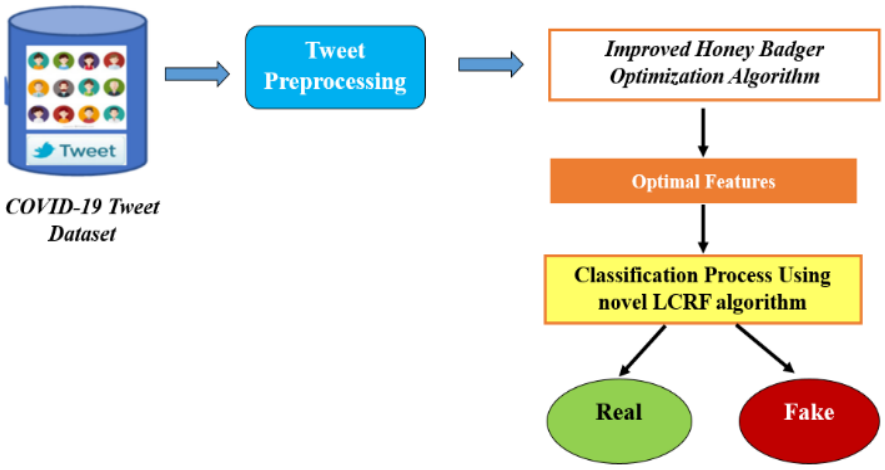

Figure 4: Proposed Architecture of LCRF-HB

Initially, the coronavirus tweet dataset is preprocessed by following the steps namely, tokenization, stop words removal, and lemmatization.

5.2.1 Improved Honey Badger Optimization Algorithm

The improved Honey Badger (HB) algorithm is developed to overcome the issue of convergence in local optimum due to population diversity loss. The proposed algorithm integrates the opposition-based optimization technique with the Honey Badger optimization algorithm to enhance the convergence in global optimum by preserving the diversity of the population during search simulations. In the Improved HB algorithm, the new solutions are obtained during the digging phase and honey phase. Then, during each iteration, the current iteration's best fitness value is compared with the previous iteration's best-fit values. If the fitness value remains the same for the p number of iterations, the opposition-based technique will be evaluated for the first k number of infeasible solutions. In this way, the optimal solution achieved till now will be saved and the infeasible solutions will be replaced by new solutions.

5.2.2 Lightweight Convolutional Random Forest Model

Then, a lightweight Convolutional Random Forest model is developed to classify the tweets as genuine or fake. A lightweight Convolutional

Neural Network is designed with a minimum computing unit node set and connecting dual edge node-set. The computing unit node-set is used for convolution operations. The connecting dual edge node-set is used for connectivity. If the computing nodes are not adjacent to each other, the edge node rejects it with a definite probability. Random Forest is generated with the n number of decision trees. The random forest technique maximizes decision tree diversity by creating many training subsets, leading to improved efficiency. The classification results obtained through each decision tree are combined based on voting to obtain the result as fake or real.

5.2.3 Experimental Results

Experimentation of the proposed model is carried out using Python and Scikit Learn library, keras, and TensorFlow packages.

The proposed LCRF-HB algorithm is compared with the existing State-of-the-art optimization feature selection methods namely, GWO, SSO, Particle Swarm Optimization (PSO), Cross-SEAN, HPO, Fire Hawk Optimizer, and Firefly optimizer for COVID-19 fake tweet detection. The accuracy comparison is depicted in Figure 5.

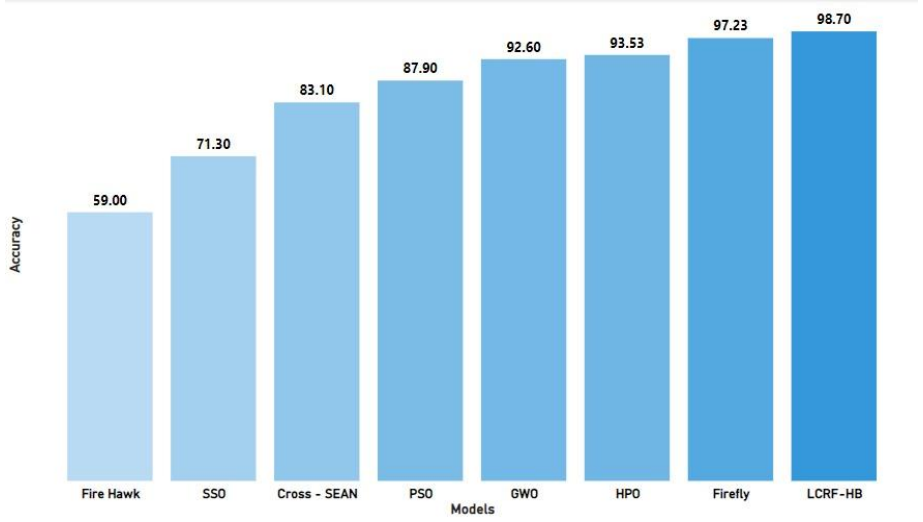

Figure 5: Performance analysis of State-of-the-art optimization based Fake news detection algorithm

The proposed algorithm achieves 98.7% accuracy and outperforms other optimization feature selection algorithms. The performance

analysis of the various state-of-the-art algorithms namely, FakeBERT, GCN, and ConvNet is compared with the proposed model and is depicted in Figure 6.

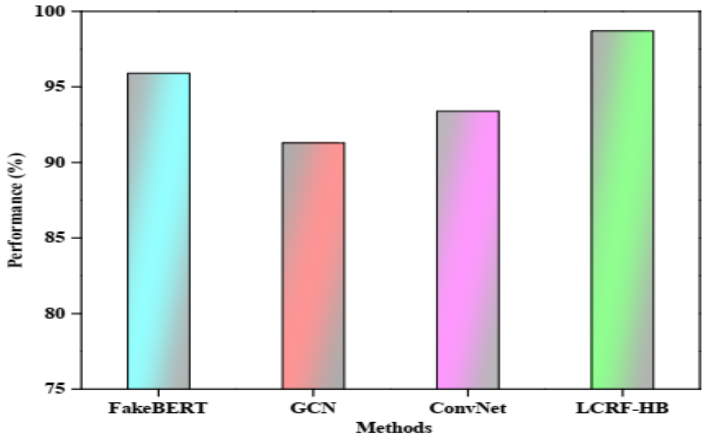

Figure 6: Performance analysis of State-of-the-art optimization based Fake news detection

The performance of FakeBERT attains 95.9%, GCN attains 91.3%, ConvNet attains 93.4% and the proposed LCRF-HB attains a greater efficiency of 98.76% and is shown in the Table 1.

Table 1: Performance analysis of State-of-the-art Fake News Detection Algorithms

Models	Accuracy
FakeBERT	95.9%
GCN	91.3%
ConvNet	93.4%
LCRF-HB	98.7%

The efficiency of training and validation accuracy concerning the epochs is depicted in Figure 7(a). The training accuracy of the proposed (LCRF-HB) model attains 80%, whereas the validation accuracy attains 70%. The effectiveness of training loss attains 0.4% and validation loss of 0.2% is depicted in Figure 7(b).

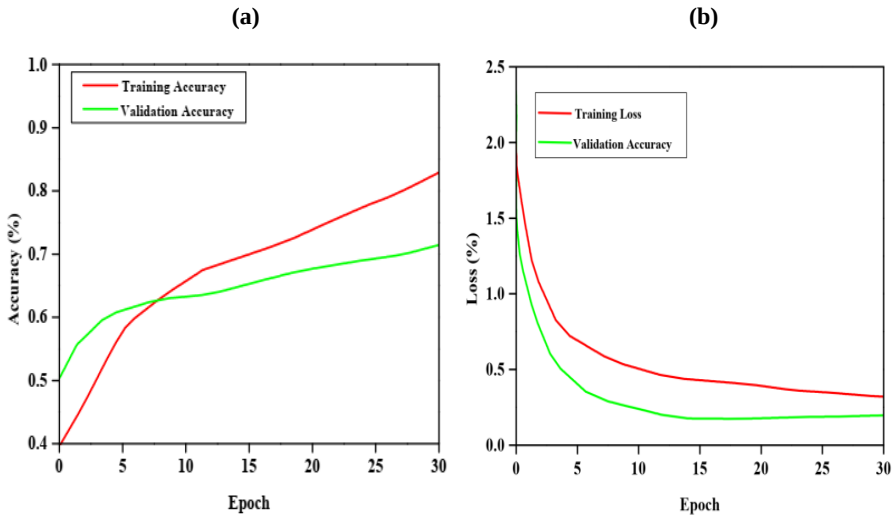

Figure 7: Efficiency of training and validation accuracy, Efficiency of training and validation loss

5.3 Credibility Detection of Emerging Tweets Using Self-trained Semi-Supervised Deep Hybrid Learning Model

The automatic COVID-19 Misinformation Detection (ACOVMD) model is proposed with the creation of an embedding layer followed by the hybridization of Lightweight Convolutional Neural Network (LCNN) and Bidirectional long short-term memory (BiLSTM), which is followed by the confidence layer and is depicted in Figure 8.

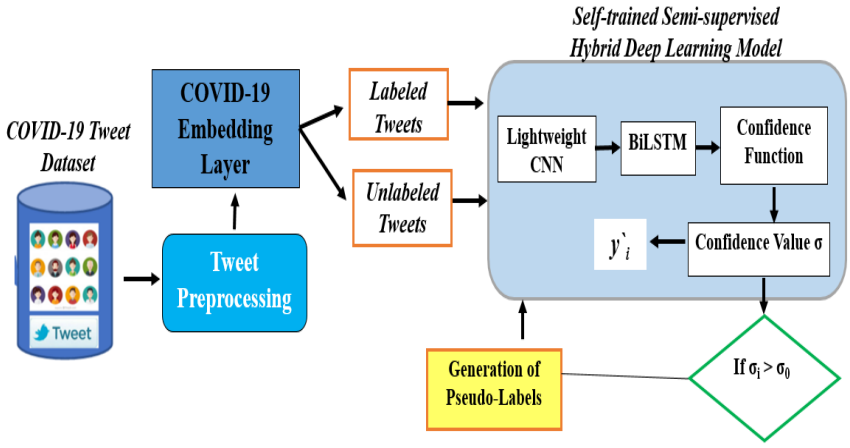

Figure 8: Proposed Architecture of ACOVMD

Initially, the coronavirus tweet dataset is preprocessed by following the steps namely, tokenization, stop words removal, and lemmatization.

The embedding layer is built to perform the embedding of the input (preprocessed tweet sequence). Each word in the tweet sequence is embedded using pre-trained COVID-19 BERT embeddings to generate contextual embeddings.

5.3.1 Training and Categorization Methodology

The proposed model trains both supervised and unsupervised components simultaneously. The supervised component having a small number of labeled tweets trains the hybrid deep learning model. Then supervised component, along with the confidence layer generates the pseudo-labels for the unlabeled tweets. Now, the generated pseudo labels and the previously available labeled tweets, form the new training set. The proposed self-trained hybrid deep learning model is made to retrain itself using this new training set. This training procedure is repeated until the pseudo labels of the previous iterations match the pseudo labels of the current iterations. Finally, the process of self-learning is done in the model and the model is updated automatically in each iteration.

5.3.2 Lightweight Convolutional Neural Network

A lightweight Convolutional Neural Network (LCNN) layer is developed in this work for the extraction of significant features from the tweets and to get rid of memory consumption issues addressed by the existing semi-supervised fake news detection models. For that, a compression method namely, separable convolution is developed to minimize the memory consumed by the neural network.

LCNN operation is performed in the following steps:

1. The contextual embedding matrix obtained from the embedding layer is multiplied with a filter in a convolutional operation to extract the matrix's most significant COVID-19 features.
2. In separable convolution, the kernel of size 3 x 3 is divided into kernel of shape 3 x 1 and 1 x 3.
3. The input embedding vectors are first convoluted with the 3 x

1 kernel and then with the 1 x 3 kernel.

4. The outputs from the preceding layer are passed through the point-wise convolution of 1 x 1 kernel size.
5. Now, the output feature maps are obtained from the pointwise convolutional layer.

5.3.3 Bidirectional Long Short-Term Memory

The Bidirectional Long Short Term Memory (BiLSTM) layer is developed to learn the long-term dependencies between the time stamps of sequential tweets and to classify the tweets as genuine or fake (credibility detection). BiLSTM is developed with two LSTMs, forward and backward LSTM.

The BiLSTM performs the following functions:

1. The output feature maps of the Lightweight Convolutional Neural Network are given as input to the forward and backward LSTM layer.
2. Using three gates, namely the Input gate, Forget gate, and Output gate, the LSTM erases some of the past COVID-19 insignificant tweet information before adding the current input
3. The outputs of the backward LSTM layer iterated from step *j* to step 1, whereas the forward LSTM layer iterated from step 1 to step *j*
4. Then, the outputs of both BiLSTM layers are concatenated and flattened into a single array of vectors by a fully connected layer.
5. Finally, this layer is followed by a sigmoid activation function to fire the output as genuine or fake.

5.3.4 Experimental Analysis

Experimentation of the proposed model is carried out using Python and Scikit Learn library, Keras, and TensorFlow packages. The proposed model split the dataset into different combinations of labeled and unlabeled data: 10% labeled tweets (5000 labeled &

45000 unlabeled), 20% labeled tweets (10000 labeled & 40000 unlabeled), 30% labeled tweets (15000 labeled & 35000 unlabeled), 40% labeled tweets (20000 labeled & 30000 unlabeled), 50% labeled tweets (25000 labeled & 25000 unlabeled), and 80% labeled tweets (40000 labeled & 10000 unlabeled). The performance of the proposed model on different combinations of labeled tweets is depicted in Table 2.

Table 2: Performance of the Proposed Model on different combinations of labeled & unlabeled Tweets

Metrics	10%	20%	30%	40%	50%	80%
Accuracy	80.9	85.2	88.8	91.9	95.5	98.1
Precision	74.4	86.4	89.3	91.8	99.0	99.5
Recall	94.1	83.6	89.1	91.9	96.1	98.3
F1 Measure	83.1	84.9	88.8	91.8	97.4	98.9

The proposed model achieves 98.1% higher accuracy for 80% labeled tweets. The training and testing accuracy of the proposed model for 80% labeled tweets is depicted in Figure 9.

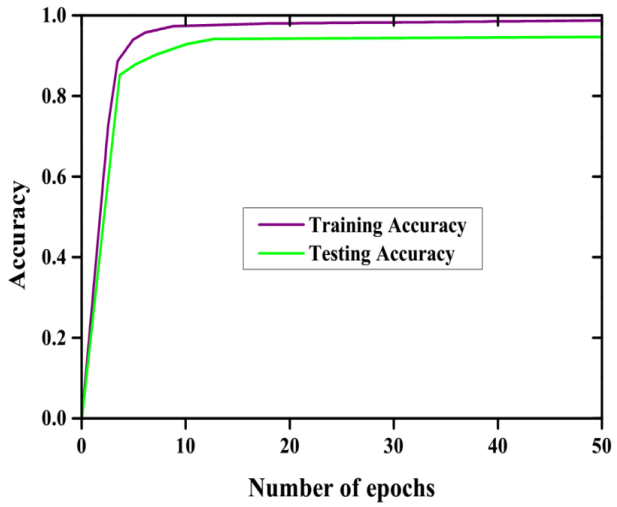

Figure 9: Training and Testing accuracy for 80% labeled experiment

The proposed model achieves 98% training accuracy and 95.6% testing accuracy for the 80% labeled experiment.

6. ORGANIZATION OF THESIS

The thesis is organized as follows:

- ***Chapter 1:*** The basic introduction to Online Social Networks and the Fake News propagating through these networks and its impact to society is discussed. The motivation for the research and the challenges prevailing in detecting fake news are discussed.
- ***Chapter 2:*** The literature survey elaborating the recent research in the fake news domain and the techniques used for detecting it with the pros and cons are discussed.
- ***Chapter 3:*** This section presents the fake news detection based on the tweet content as with the source credibility by checking the deviations from reliable sources.
- ***Chapter 4:*** This section presents the optimal feature selection based fake news detection to solve the low converge rate and population diversity loss issue by developing a Lightweight Convolutional Neural Network Honey Badger optimization algorithm.
- ***Chapter 5:*** This section presents the credibility detection of emerging tweets by developing a self-trained semi-supervised Hybrid Deep Learning model.
- ***Chapter 6:*** This section discusses the conclusion and Future work regarding the contributions of the research and future research directions.

REFERENCES

[1] N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, & D. Lazer. Fake news on Twitter during the 2016 US presidential election. *Science*, ***363***(6425) (2019), 374-378.

[2] S. Biradar, S. Saumya, & A. Chauhan. Combating the infodemic: COVID-19 induced fake news recognition in social media networks. *Complex & Intelligent Systems*, ***9***(3) (2023), 2879-2891.

[3] M. Haman. The use of Twitter by state leaders and its impact on the public during the COVID-19 pandemic. *Heliyon*, ***6***(11) (2020).

[4] C. Zhang, A. Gupta, X. Qin, & Y. Zhou. A computational approach for real-time detection of fake news. *Expert Systems with Applications*, ***221***(2023), 119656.

[5] K. P. Choi, E. H. H. Kam, X. T. Tong, & W. K. Wong. Appropriate noise addition to metaheuristic algorithms can enhance their performance. *Scientific Reports*, ***13***(1) (2023), 5291.

[6] X. Dong, U. Victor, S. Chowdhury, & L. Qian. Deep two-path semi-supervised learning for fake news detection. *arXiv preprint arXiv:1906.05659,* (2019).

[7] Z. Ghadiri,M. Ranjbar, F. Ghanbarnejad, & S. Raeisi. Automated Fake News Detection using cross-checking with reliable sources. arXiv:2201.00083, (2022).

[8] M. N. Nikiforos, S. Vergis, A. Stylidou, N. Augoustis, K. L. Kermanidis, & M. Maragoudakis. Fake news detection regarding the Hong Kong events from tweets. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (2020), 177-186. Cham: Springer International Publishing.

[9] K. A. Qureshi, R. A. S. Malick, M. Sabih, & H. Cherifi. Complex network and source inspired COVID-19 fake news classification on Twitter. *IEEE Access*, ***9***(2021), 139636-139656.

[10] N. Capuano, G. Fenza, V. Loia, & F. D. Nota. Content Based Fake News Detection with machine and deep learning: a systematic review. Neurocomputing, (2023).

[11] F.A. Ozbay, & B. Alatas. A novel approach for detection of fake news on social media using metaheuristic optimization algorithms. *Elektronika ir Elektrotechnika*, ***25***(4) (2019), 62-67.

[12] M. Abd Elaziz, A. Dahou, D. A. Orabi, S. Alshathri, E. M. Soliman, & A. A. Ewees. A Hybrid Multitask Learning Framework with a Fire Hawk Optimizer for Arabic Fake News Detection. *Mathematics*, ***11***(2) (2023), 258.

[13] A. Jaiswal, H. Verma, N. Sachdeva, A. Wary, G. Indra, &

D. K. Brahma. A Meta-Heuristic Based Approach for Optimal Fake News Detection Using Supervised Learning (2023).

[14] H. J. Alshahrani, A. Q. Hassan, K. Tarmissi, A. S. Mehanna, A. Motwakel, I. Yaseen, & M. I. Eldesouki. Hunter Prey Optimization with Hybrid Deep Learning for Fake News Detection on Arabic Corpus. *Computers, Materials & Continua*, ***75***(2) (2023).

[15] S. H. Nithya, & A. Sahayadhas. Meta-heuristic Searched-Ensemble Learning for fake news detection with optimal weighted feature selection approach. *Data & Knowledge Engineering*, ***144***(2023), 102124.

[16] P. Patwa, S. Sharma, S. Pykl, V. Guptha, G. Kumari, M. S. Akhtar, & T. Chakraborty. Fighting an infodemic: Covid-19 fake news dataset. In *Combating Online Hostile Posts in Regional Languages during Emergency Situation: First International Workshop, **1*** (2021), 21-29. Springer International Publishing.

[17] M. K. Elhadad, K. F. Li, & F. Gebali. Detecting misleading information on COVID-19. *IEEE Access*, ***8***(2020) 165201-165215.

[18] W. S. Paka, R. Bansal, A. Kaushik, S. Sengupta, & T. Chakraborty. Cross-SEAN: A cross-stitch semi-supervised neural attention model for COVID-19 fake news detection. *Applied Soft Computing*, ***107***(2021), 107393.

LIST OF PUBLICATIONS

Journal Publications - SCI

1. Birunda, S. S., Devi, R. K., Muthukannan, M. "An Efficient Model for Detecting Covid-19 Fake news using Optimal Lightweight Convolutional Random Forest. *Signal, Image and Video Processing*. (Springer Publisher)

Journal Publications – Scopus

2. Birunda, S. S., Devi, R. K., Muthukannan, M., & Babu, M. M. ACOVMD: Automatic COVID-19 misinformation detection in Twitter using self-trained semi-supervised hybrid deep learning model. *International Social Science Journal*. (Wiley Publisher)

3. Shanmugavel, S. B., Devi Rangaswamy, K., & Muthukannan, M. (2023). A Deep Learning Model to Detect Fake News about COVID-19. *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), 16*(9), 58-66. (Bentham Science Publisher)

4. Selva Birunda, S., Kanniga Devi, R., Suresh Kumar, N., & Muthukannan, M. On the Effect of Intelligent Neurally Processed Features on COVID-19 Fake News Detection Performance, Artificial Intelligence in the Life Sciences, (Elsevier Publisher)

Conference Publications

5. Selva Birunda, S., & Kanniga Devi, R. (2021). A review on word embedding techniques for text classification. *Innovative Data Communication Technologies and Application: Proceedings of ICIDCA 2020*, 267-281. (Springer Publisher)

6. Birunda, S. S., & Devi, R. K. (2021, March). A novel score-based multi-source fake news detection using gradient boosting algorithm. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)* (pp. 406-414). IEEE.